EVALUATING OUR WEB PRESENCE:

CHALLENGES, METRICS, RESULTS

A Strategic Approach to Web Evaluation

Fred B. Wood

Special Expert, National Library of Medicine

Office of Health Information Programs Development

Chair, CENDI Task Group on Web Metrics & Evaluation

Tuesday, April 17, 2001

CENDI

UNITED STATES National Library of Medicine

In Native American tradition, the eagle stands for*

--vision

--inspiration

--clarity

--wisdom

--new beginnings

May these be our symposium goals re web evaluation.

(*from Sun Bear, Wabun Wind, and Crysalis Mulligan, *Dancing With the Wheel: Medicine Wheel Workbook*, Fireside/Simon & Schuster, 1991, pp. 173, 179)

Outline

- Overview--Why this symposium?
- CENDI Task Group on Web Metrics & Evaluation
- NLM Work Group on Web Evaluation
- Illustrative Evaluation Results
- Internet Audience Measurement
  - Health Information Sector
  - Government Sector
- Online User Survey--MEDLINE*plus*
- Internet Connectivity Performance

The Web Evaluation Challenge

- Much more difficult to assess outcomes in the Internet/Web world.
- Much harder to get meaningful, direct feedback from users (visitors, consumers).
- The “system” is more complex, with many more players and options.
- Traditional “evaluation” methods are less applicable or feasible.
- New or re-invented approaches are needed.

(W. Ross) Ashby's Law of Requisite Variety*

Complexity of Monitoring/Evaluation and Management System must ≥ Complexity of Real World System

--applies in spades to the web-based information realms of many government and private sector organizations

*W. Ross Ashby, *Introduction to Cybernetics*, Routledge, 1964. Also see references in Stafford Beer, *Decision and Control*, John Wiley, 1966.

Meeting the Challenge--Covering the Bases

- Speakers will cover the bases pretty well
- Fred Wood, overview + online panel, survey, Internet testing, cross comparisons
- Kevin Mabley--online surveys, competitive analyses
- Kevin Garton--audience measurement w/ online panels
- Bill Silberg--cross-cut from industry perspective
- Carlynn Thompson--cross-cut from government agency perspective
- Lisa Zolly--usability testing
- Susan Elster--focus groups, interviews, field surveys

CENDI

UNITED STATES National Library of Medicine

CENDI Web Metrics Task Group

--Web Metrics & Evaluation identified as a priority area
--1998 initial CENDI study conducted
--2000 study intended to provide update re:
agency activities & identify action priorities
--2000 work group members included reps from 7
CENDI agencies
--1998 and 2000 studies available on CENDI web site @ www.dtic.mil/cendi

CENDI Study Results in Brief

--all agencies use web log analysis software
--metrics tracked vary but typically include
   hits, page views or pages downloaded,
   unique visitors, number of user sessions
--online web user surveys conducted although with
   mixed results
--usability studies typically a part of web design
--most use some form of Internet connectivity monitoring but details vary

CENDI Study Outcome

--NEED common web metrics framework
--NEED CENDI web watch--standard core of web
metrics/definitions, best metrics practices
--NEED CENDI Internet watch--standard core of
Internet performance metrics/definitions,
best practices
--NEED better mechanisms for sharing
knowledge and experience, e.g., training,

NLM Web Evaluation Work Group

- Web Metrics & Evaluation identified as a priority area
- NLM Director appointed agency-wide work group
- Mandated to explore and pilot test a range of evaluation methods
- Reviewed related studies on web evaluation methods and examined public/private sector options
- First wave of pilot tests almost complete

NLM Work Group Activities

- Tested and deployed a wider variety of evaluation methods appropriate to web-based environment
- Tried external web usage monitoring capabilities
- Tried randomized on-line user (visitor) survey
- Integrated new methods with traditional methods, e.g., survey research for benchmarking, internal web metrics
- Contracted for comparative web site analysis
- Implemented Internet/web connectivity testing

Internet Audience Measurement Online Panels

- NLM contracted with PCDataOnline, Inc. for external usage data based on PCData's online panel of 120K home users who agreed to monitoring of home web usage
- Results extrapolated to US home Internet market
- PCData now out of business--client base bought by comScore Networks, Inc.
- Kevin Garton of comScore will address methods, services, and issues

Illustrative Audience Measurement Data US Home Internet Market--Top Level Domains

- As of February 2001 (last month of PCData data)
- Health information web sites--*webmd.com*, 6.6M unique users; *allhealth.com* (iVillage), 3.4M; *ediets.com*, 2.9M; *nih.gov*, 2.5M
- US Government web sites--*irs.gov*, 7.4M unique users; *usps.com*, 4.4M; *fedworld.gov*, 3.9M; *ustreas.gov*, 3.9M; *nasa.gov*, 3.5M; *ed.gov*, 3.3M, *nih.gov*, 2.5M
- Before tax and student aid seasons (Nov 2000)--*usps.gov*, 3.8M; *nasa.gov*, 2.3M; *nih.gov*, 2.3M

Illustrative Audience Measurement Data US Home Internet Market--NIH Drill Down

US Unique Home Users (% of NIH Total) [source: PCData]

	Dec 2000	Jan 2001	Feb 2001
Domains			
nlm.nih.gov	42.6%	45.3%	50.4%
niddk.nih.gov	11.2	8.7	7.9
nci.nih.gov	7.1	7.4	8.7
nhlbi.nih.gov	4.4	3.8	6.0
Web sites			
MedlinePlus	14.2%	18.0%	17.5%
PubMed	13.4	18.3	20.1
www.nih.gov	13.6	11.1	17.0
www.niddk.gov		11.2	8.7 7.9

Online User Survey--MEDLINE*plus*

- NLM contracted with CyberDialogue, Inc., for development and implementation of a randomized online survey of MEDLINE*plus* users
- Survey instrument was developed collaboratively by NLM and CyberDialogue, and cleared under NLM's blanket OMB approval process.
- Sample size ~3,000 users over ~ 10 days in February 2001.
- Kevin Mabley of CyberDialogue will provide details on the methods and illustrative results.

Illustrative Cross-Comparisons--MEDLINE*plus*

- High site loyalty--42% visit M+ monthly or more often, compares to 4-18% range for top 5 commercial health sites (based on Cyber Dialogue's CybercitizenHealth nationwide phone survey).
- High home use--61% search from home, similar to Cybercitizen Health and PCData survey results.
- High focus on specific disease, condition, diagnosis--63%, again similar to CybercitizenHealth and PCData surveys.

Cross-Comparison with Web Log Data--MEDLINE*plus*

- Time-of-day/day-of-week variability--97% correlation between survey data and web log data.
- US/non-US user split--73% US/27% non-US per survey, 65% US/35% non-US per web log data (but majority of log data cannot be resolved).
- Repeat visitors (2+ times per month)--36% per survey, 21% per web log data (but log data subject to error factors both ways)
- Survey non-response bias may be relatively minimal.

Evaluating End-to-End Internet Connectivity

- How fast can users download web pages and conduct database searches via the Internet?

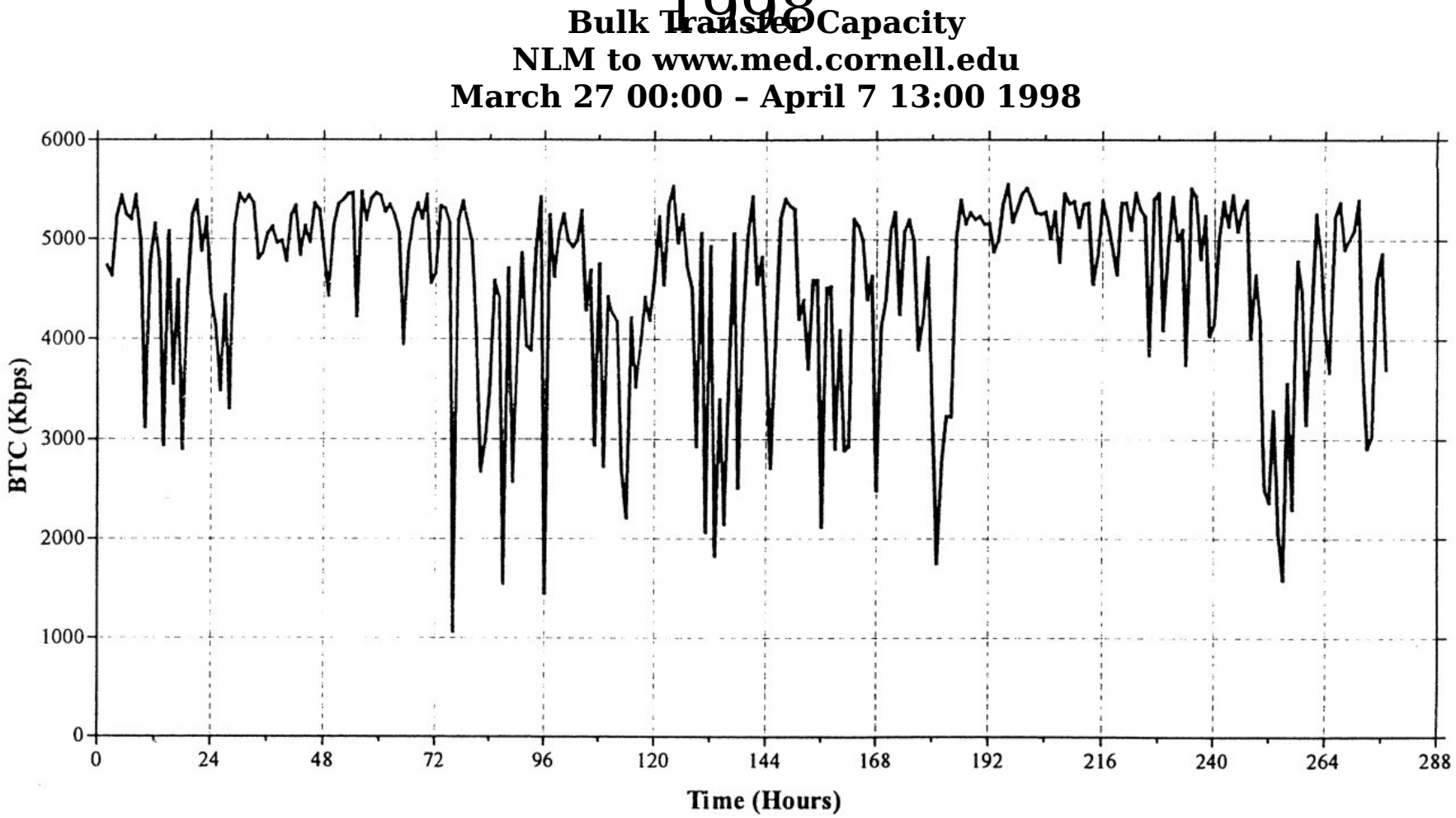
- Metrics--transport level/TCP (BTC, RTT, route stability, packet loss), applications level/HTTP (response time, download time)
- Timing--ad hoc, defined period, continuous
- Test nodes--user terminals, web sites, network servers
- Test direction--asymmetric/symmetric, inbound/outbound
- Test network--custom, benchmark; selected by providers, users, and/or commercial services

Comparison of Internet Connectivity Test Results Selected Pathways NLM to Host Web Sites April 1998 and April 2001 (courtesy V. Cid)

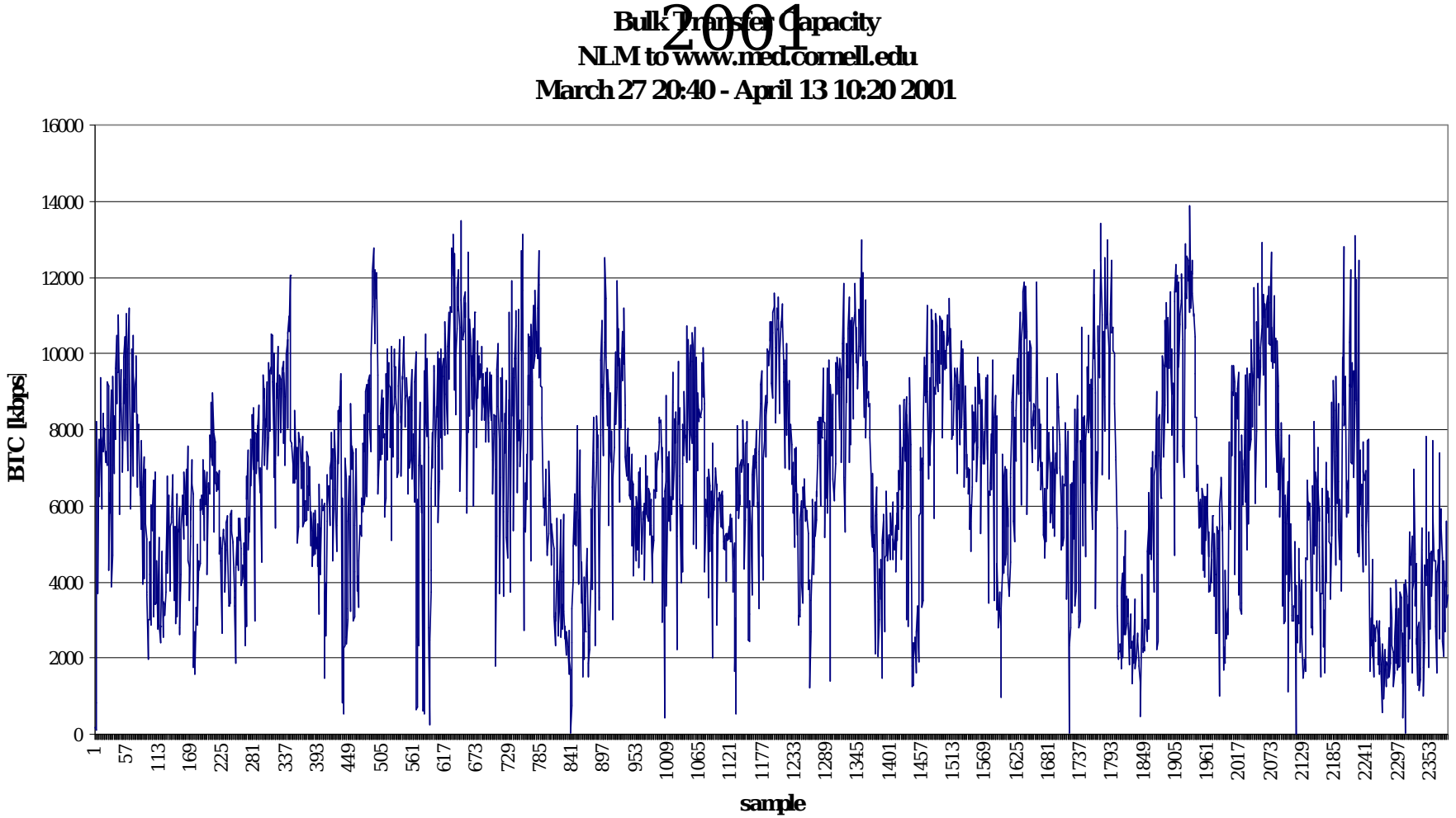
Host name	Median BTC [Mbps]		Mean RTT [ms]	
	1998	2001	1998	2001
www.med.cornell.edu	4.91	6.85	18	14.59
www.library.ucla.edu (*)	2.67	3.96	70	72.81
www.uicomp.uic.edu (*)	0.16	1.32	65	40.48
www.hslib.washington.edu (*)	3.74	5.16	83	74.44
www.cisti.nrc.ca (*)	0.53	8.02	84	53.26
www.dimdi.de	0.12	13.33	160	116.23

(*) Paths between NLM and these hosts are today (2001) through the Abilene network

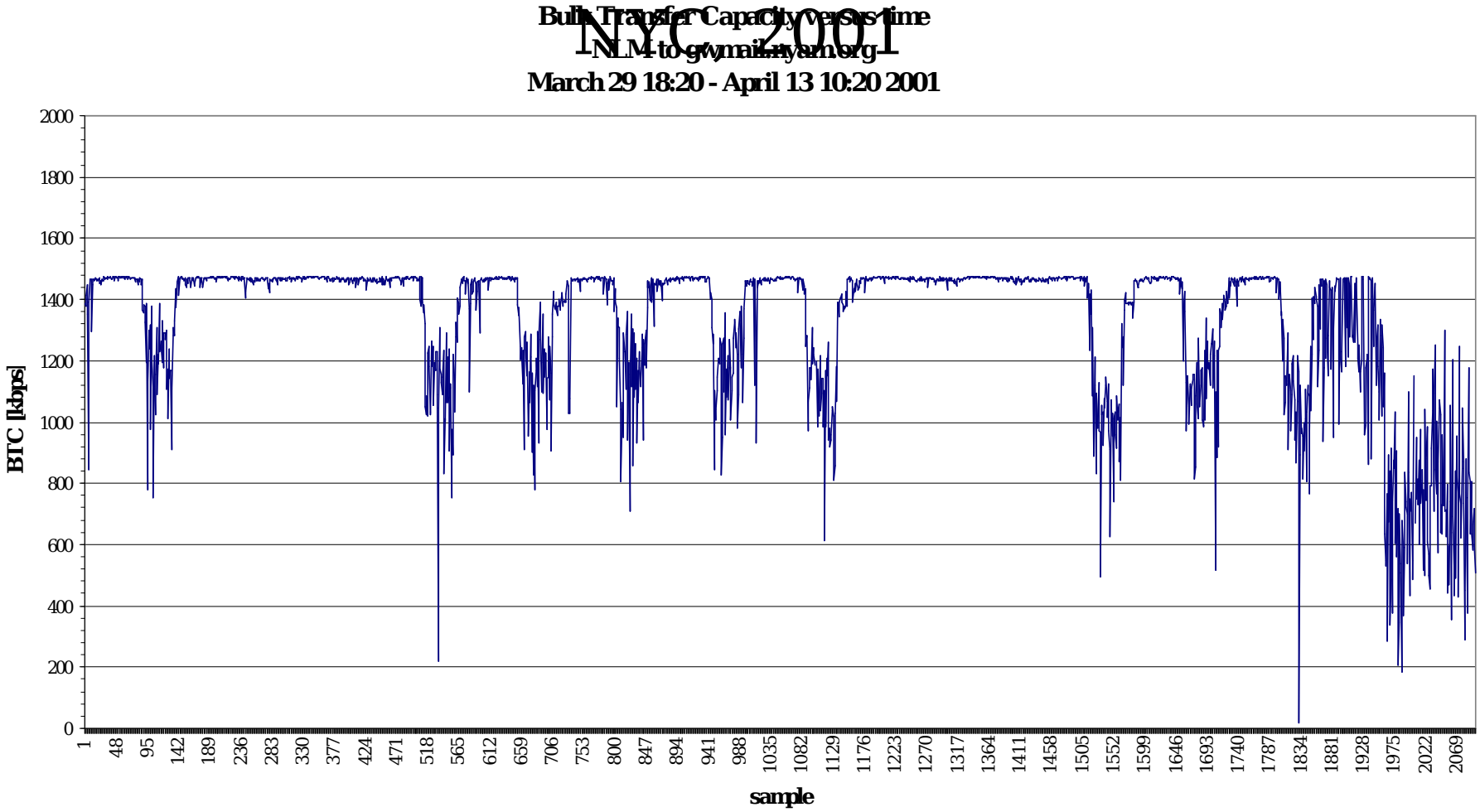
Illustrative Internet Connectivity Test Results-- NLM to Cornell Medical College, NYC, 1998



Illustrative Internet Connectivity Test Results-- NLM to Cornell Medical College, NYC, 2001

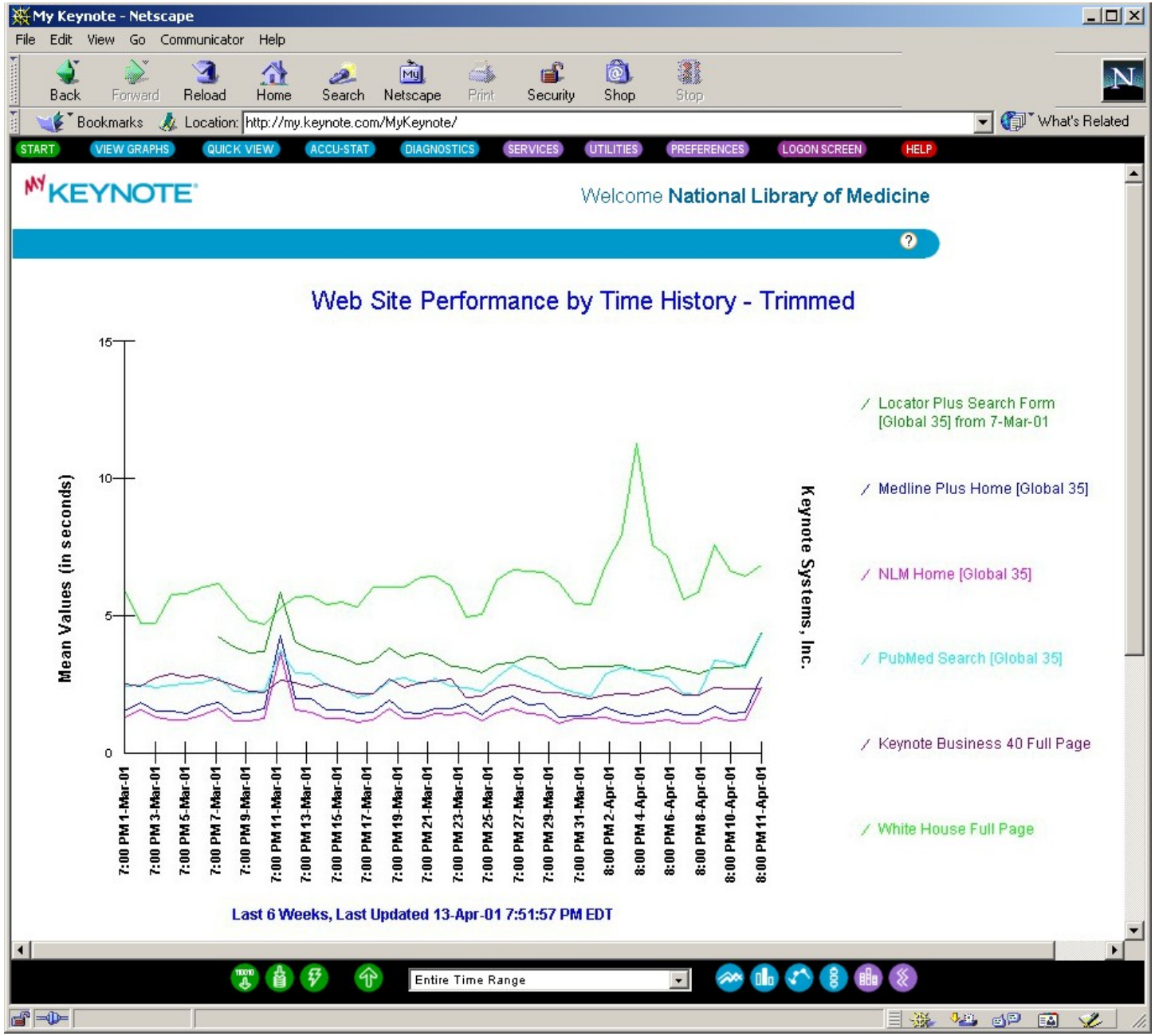


Illustrative Internet Connectivity Test Results-- NLM to New York Academy of Medicine, NYC2001



Illustrative External Internet Connectivity Testing
Keynote, Inc., Full Page Web Download, Mar/Apr 2001

- Measures total time (in seconds) to download full web page from proxy user locations in the U.S. and abroad
- Measures component time segments, e.g., DNS time (to resolve DNS to IP address), connection setup time, time to first byte, redirect time, content time--for diagnostics.
- Current proxy locations include 25 major U.S. cities and up to 25 major foreign cities (w/ T-1+, Sun machines).
- Useful for comparing web performance against benchmark sites and groups of sites, and for isolating anomalies.
- Keynote Business 40 = average web performance of 40 major business web sites.


My Keynote - Netscape
File Edit View Go Communicator Help
Back Forward Reload Home Search Netscape Print Security Shop Stop
Bookmarks Location: http://my.keynote.com/MyKeynote/
What's Related
START VIEW GRAPHS QUICK VIEW ACCU-STAT DIAGNOSTICS SERVICES UTILITIES PREFERENCES LOGON SCREEN HELP
MY KEYNOTE
Welcome National Library of Medicine
Web Site Performance by Time History - Trimmed
Mean Values (in seconds)
15
10
5
0
7:00 PM 1-Mar-01
7:00 PM 3-Mar-01
7:00 PM 5-Mar-01
7:00 PM 7-Mar-01
7:00 PM 9-Mar-01
7:00 PM 11-Mar-01
7:00 PM 13-Mar-01
7:00 PM 15-Mar-01
7:00 PM 17-Mar-01
7:00 PM 19-Mar-01
7:00 PM 21-Mar-01
7:00 PM 23-Mar-01
7:00 PM 25-Mar-01
7:00 PM 27-Mar-01
7:00 PM 29-Mar-01
7:00 PM 31-Mar-01
8:00 PM 2-Apr-01
8:00 PM 4-Apr-01
8:00 PM 6-Apr-01
8:00 PM 8-Apr-01
8:00 PM 10-Apr-01
8:00 PM 11-Apr-01
Keynote Systems, Inc.
Locator Plus Search Form [Global 35] from 7-Mar-01
Medline Plus Home [Global 35]
NLM Home [Global 35]
PubMed Search [Global 35]
Keynote Business 40 Full Page
White House Full Page
Last 6 Weeks, Last Updated 13-Apr-01 7:51:57 PM EDT
Entire Time Range

Why evaluation matters--desired outcomes for NLM (I)

1. Increase awareness of health information
2. Increase accessibility of health information
3. Increase usage of health information
4. IMPACTS of usage of health information
5. + Impact on patient/provider decisions
6. + Impact on individual and overall health
7. Reduce health disparities (among ethnic, cultural, geographic groups)

UNITED STATES
National Library of Medicine

Why evaluation matters--desired outcomes for NLM (II)

8. Reduce digital divide (computer/Internet haves and have nots, moving target, includes training)
9. Improve quality of health information
10. Improve quality of health information professionals and intermediaries
11. Encourage innovation in medical informatics, telemedicine, Health Next Generation Internet
12. Encourage successful application of innovations in health arena

In Conclusion

Evolution in web evaluation? For sure.

Revolution? Probably.

What kind of revolution?

We need innovative web systems evaluation to help chart our course.

Science is progress.

UNITED STATES
National Library of Medicine